Web search for translation: an exploratory study on six Chinese trainee translators' behaviour

Translators' information skills have long been recognised as one of the most important competence for translators these days (e.g. Gouadec 2007, Pym 2011, Garcia 2009, Mikhailov 2015). Of particular interest is the perception of the web as an invaluable tool for locating and confirming terms for specialised translation (Harman 2002). The intention for this paper is to investigate Chinese trainee translators' web-search behaviour in this context. Screen recording was supplemented by think-aloud protocols as data collection methods. The data suggest that Chinese trainee translators use a variety of web resources, which largely tallies with previous researchers' findings (e.g. Massey and Ehrensberger-Dow 2011, Daems et al. 2016). Two types of web-search behaviours are identified. In addition, interesting idiosyncratic behaviours are found in terms of how and potentially why different types of web resources are used in conjunction with each other by individual trainees. Relevant pedagogical implications are drawn as a result.

Keywords: web search, screen recording, information literacy, translation pedagogy, cognitive behaviour.

1. Introduction

In the volume, *Describing Cognitive Processes in Translation*, Ehrensberger-Dow et al (2015) address translation as an act and event. On the one hand, translation as an act reflects a more traditional focus of research on translators' mental processes, whereby translators were largely perceived as individuals who work in isolation. On the other hand, translation as an event reflects a more recent line of research, whereby translators are regarded as working in a wider social context, interacting with other agents and with technology. These two lines of interest in translation process research are largely consistent with Chesterman's (2009) proposition of "Translator Studies" consisting of cultural, cognitive and sociological strands, as opposed to "Translation Studies". The present study can be positioned in the cognitive and potentially sociological strands of "Translator Studies", since it taps into the field of human and computer

interaction, whereby web-search behaviours are observable acts and events at the interface between the translator as an individual and the wider information source, the World Wide Web.

1.1 Web Searches in Translation

Recognised as a key translation competence, translators' use of documentation and information sources has long been investigated in translation process research by scholars (e.g. Krings 1986; Jääskeläinen 1989; Künzli 2001; Livbjerg & Mees 2003; Wakabayashi 2003; Zheng 2014). However, until recently, much translation process research focus on paper-based dictionaries and printed resources; yet it is hard to ignore that most translators nowadays rely heavily on digital resources rather than paper-based ones. (Mikhailov 2015)

In particular, the World Wide Web has been recognised as a vital source of information, both in the profession and in the academia. On the one hand, specialised glossaries and online resources are commonly shared and exchanged on the web by professional translators. For example, Mayor Serrano (2009) posts specialised online resources and glossaries on the web for Spanish/English professional translators working in the field of marine and naval architecture. In the field of medical translation, Maniez (2007) demonstrates how the web can be used as a corpus for verifying medical terms. On the other hand, in the academia, web-based information skills have been increasingly recognised as an important aspect of translator training. Pinto & Sales (2007, 2008a, 2008b) advocate the importance of information literacy. In particular, in their questionnaire studies, it was found that the single most important information need for translator training is 'information search' (Pinto & Sales 2008a, 60). Zanettin (2002) describes how the web can be used to compile mini-corpora for translator training. In a similar vein, Buendía-Castro & López Rodríguez (2013) highlight that in addition to compiling a corpus, the web can also be used as a corpus for translators. However, there are very few empirical studies specifically tapping into web search as an important aspect of the translation process in translation studies.

Enríquez Raído (2011, 2014) is one of the first scholars who has specifically focused on web searches by translators. She used a combination of pre-task questionnaires, screen recording,

online search reports and post-task interviews to collect her data. Her four subjects were primarily students from her scientific and technical translation module, even though she also reported on two professionals in her pilot study[i]. Her translation students were asked to complete two translation tasks. Task 1 was released and completed by her students in week 3, whereas Task 2 was released in week 4 of the module, which lasted for 12 weeks. Task 2 was designed to be more challenging than Task 1. In fact, most of her results and findings could be attributed to this very design. For instance, she found that her student subjects tended to visit the same web resources repeatedly, although in Task 2, as opposed to Task 1, they also appeared to use more varied web resources, probably due to the greater complexity of Task 2. Furthermore, in Task 2, her subjects developed a "more interactionistic style of online search", whereas in Task 1, her subjects demonstrated "a shallow online search style" (Enríquez Raído 2011, 70), indicating that in Task 1 her subjects mainly used bilingual or monolingual online dictionaries to look for translation equivalents or to solve comprehension problems, whereas in Task 2 they began to engage with other online resources in order to find more background information. In addition, she found that her student subjects tended to start their web-search process with online dictionaries. Only when online dictionaries failed to provide satisfactory equivalents did they start to use other web resources, such as search engines.

Enríquez Raído's project was embedded in translation didactics. For example, she asked her students to write online search reports, which involve students' self-evaluation of their web-search behaviours during and/or after their web search. Although these online search reports offer great didactical value, they also have some drawback because these online search reports may interrupt natural flows of the web-search process. Enríquez Raído's (2011, 337) proposed a model for the cognitive process of translators' web search, starting with 1) the formulation of a search need, followed by 2) the specification of a search need, 3) the implementation of a search process, and finally 4) the selection of a search outcome.

Similar to Enríquez Raído's study, the present study taps into trainee translators' web search behaviours. However, instead of covering the web-search process as a whole, it focuses on what Enríquez Raído's called, "implementation of a search process" and aims to depict qualitatively how, what and potentially why web searches are conducted by trainee translators particularly in the language direction of English to Chinese. To the best of the author's knowledge, this

study is the first to examine trainee translators' web-search behaviour in this language combination. The following list of operational questions encapsulates the general aims of the present study.

1. What and how web resources are used by Chinese trainee translators?
2. What search queries[ii] are made by Chinese trainee translators?
3. What and why idiosyncratic web-search behaviour is adopted by individual Chinese trainee translators?
4. What potential pedagogical implications can be drawn from the Chinese trainee translators' web-search behaviour?

2. Data collection

This research project adopts two data-collection methods: screen recording supplemented by think-aloud protocols. Screen-recording software is particularly suited to the present research project because it does not just record translators' web-search behaviours on their computer screens, but also records their think-aloud protocols in real time. The combination of these two research methods is particularly effective, because it captures not only comprehensive records of translators' web-search activities but also the verbalisation of their own behaviours simultaneously.

Six Chinese trainee translators[iii] from the MA in Translation Studies at the University of Leicester, U.K., were invited to participate in the present study. Their consent was sought via a form which detailed data anonymity, non-interference with their academic attainment, etc. It is important to note that these subjects were not specifically trained for web search in their curriculum even though it cannot be ruled out that some of them may have been better equipped in terms of their information skills prior to their postgraduate study in the U.K. They were instructed to conduct web searches for the purpose of translating a piece of scientific/technical text, entitled "Viking builds world's first large LNG passenger ferry"[iv]. The source text was about 500 words in length. It was an excerpt from a newsletter, *Horizon: News and Information for the Marine Industry, A Lloyd's Register Magazine*, May 2011 issue. The source text was chosen because, even though its content belongs to a specialised subject area and contains many specialised terminology, it was written in simple and plain English that is unlikely to cause

confusion. It is worth pointing out that given the nature of the source text, an original assumption was that postgraduate trainee translators who were unfamiliar with this specialised domain (i.e. naval architecture) were likely to search the web prior to starting their translation per se. In other words, it was logical to assume that trainees would focus on this pre-translation phase. However, this turned out to be not the case for all subjects, as among the six subjects, four of them conduct web search prior to the translation per se and two of them only search the web while translating.

Subjects were shown how to download and use BBFlashBack Express, a piece of screen-recording software (http://www.bbsoftware.co.uk/), on their own PCs in advance of the task. In addition, they were first trained to think aloud in class and then on separate occasions asked to practise using BBFlashBack Express to record their on-screen translation behaviour and thinking aloud (by enabling a 'Record Sound' function in BBFlashBack Express) . The actual tasks themselves, however, were conducted in trainee translators' own habitual working environment (not in classrooms) using their own PCs and in their own time[v] without any intervention from the researcher. This research design was aiming to mimic the authentic conditions of Chinese trainee translators who work independently on their translation assignments outside the classroom environment. The advantage of this lies in its maximum ecological validity where trainees' authentic idiosyncratic behaviours can be observed. The disadvantage of this design, however, lies in the fact that the researcher was not able to monitor and control the experiments in real time should any technical or other unexpected issues occur.

The screen-recording data were largely transcribed and coded based on Enríquez Raído's (2014, 103-105) coding of screen-recording data. They include: time frame, window/URL, query, and TAP/Note. "Time Frame" indicates approximate time in relation to the on-screen behaviours being recorded. "Window" indicates a dominant window or tab that was open on the screen. "URL" indicates a website or webpage on the screen. TAP/Note indicates think-aloud protocol transcriptions and relevant notes (based on the on-screen data) in relation to the timeframe. "Query" is an important data source that was recorded in the present study and is directly related to the research question, namely "What search queries are made?", which warrants further explanation. In many ways, queries can be considered to be a manifested component of Question 1, i.e. what and how web resources are used. Queries are also one of the most overt

and explicit features of web search and hence deserve to be looked into. In fact, it is important firstly to distinguish between the concepts of "query" and "term". A mechanical definition of "query" is "a string of zero [and/or] more characters entered into the Web Information Retrieval system" (Jansen & Pooch 2001, 244; cf. Enríquez Raído 2011, 374), whereas "term" is defined as "a string of characters separated by some delimiter such as space, a colon, or a period" (ibid). In other words, a query can contain a number of terms, symbols (such as Boolean operators, e.g. +, -), and/or many other characters. It is worth adding that, for the purposes of the present study, the working definition of "query" has been extended to include queries entered outside search engines (e.g. online dictionaries) since the present study aims to investigate Chinese trainee translators' web-search behaviours rather than their use of search engines exclusively.

3. Results and Discussion

3.1 What and how web-based resources were used

In terms of the varieties of web-based resources used, findings from the present study are largely comparable to those emerging from Massey and Ehrensberger-Dow's large-scale survey of 110 Swiss translators (2011, 4) and Daems et al's study of trainee translators' translation and post-editing process (2016). To sum up, four types of web resource were used by the Chinese trainee translators in the present study, i.e. search engines, online dictionaries, online machine translation tools, online encyclopaedias. These are shown in Table 1 below.

INSERT TABLE 1 HERE

In terms of the search engines, perhaps not surprisingly, Baidu (being one of the most popular Chinese search engines) and Google were the search engines of choice by the Chinese trainee translators. This finding is in line with Xu and Wang's (2011, 73-74) questionnaire study. In terms of dictionaries, the subjects of the present study used either one or a combination of the following three dictionaries: ICIBA, Youdao, and Lingoes. While ICIBA and Youdao are bilingual (English/Chinese) dictionaries, Lingoes is a multilingual one. However, there is a distinctive difference between Lingoes and the other two online dictionaries, because Lingoes

is a customised suite of dictionary package that has to be downloaded first. According to Lingoes's website, "Lingoes is a new-generation dictionary and translation software." (http://lingoes.net/en/translator/index.html) Effectively, this is a free e-package that incorporates various online dictionaries, statistics-based machine translation (SMT) software (e.g. Google Translate) and online encyclopaedia (e.g. Wikipedia). Users of Lingoes can choose which online dictionaries, MTs, and encyclopaedia, to install. Once customised and installed, Lingoes automatically searches all the web resources (provided that the PC has internet access) and displays its search results in a separate drop-down window when prompted by placing a cursor on any words, phrases or sentences appearing in Microsoft Word or a web browser, e.g. Internet Explorer. An example of this is shown in the screenshot below.

INSERT FIGURE 1 HERE

Figure 1 Screenshot of Subject F's use of Lingoes

As seen in Figure 1, Subject F selected the word, "sought" and placed her cursor on the word. Then, a window was opened automatically showing relevant searched results. In this instance, there were three: one from an English-Chinese and the other two from English-English dictionaries. The fact that Subject F was able to place her cursor and check multiple online dictionaries and web resources easily and simultaneously, appear to dramatically alter the dynamics and even ergonomics of how she used web resources, compared with other trainees in the study. It is worth noting here that Subject F was one of the two trainees who chose to skip the pre-translation web search phase. Subject F's web search behaviours will be discussed in details in Section 3.3.2.

Apart from Subject F, Subject D also chose to use machine translation software, named, "Youdao Fanyi". Similar to Google Translate, Youdao Fanyi is SMT. What is interesting about Subject D is that she did not just choose to use Youdao Fanyi as her main web resource, but also used it as if it were an online dictionary. Again, Subject D's idiosyncratic behaviour will be discussed in more detail in Section 3.3.1.

One interesting observation about the use of web resources is that three of the six subjects (i.e. Subject A, Subject B and Subject C) predominantly relied on online dictionaries more than any other types of web resources (as seen in Table 2). This coincides with many of the early studies in translation process research, where bilingual dictionaries are often the first port of call and dominant information sources for novice/trainee translators. From this perspective, the present study shows that even though trainees may no longer rely on paper-based dictionaries, for some, bilingual online dictionaries are still their primary sources of information. Again, this coincides with Enríquez Raído's (2014) finding of her students' web-search behaviour, as mentioned in Section 1.1.

Table 2 provides an overview of the numbers of times that each web resource was consulted by individual trainee translators.

INSERT TABLE 2 HERE

Based on the data presented in Table 2, one can see that regardless of how many uses of web search were made in total by each subject[vi], there appear to be two types of web search styles. One type of trainees' web-search behaviour can be characterised by what might be called, 'single web resource approach'. In other words, only one web resource was used extensively. For example, Subject A preferred to use ICIBA, Subject C Youdao, Subject D Youdao Fanyi and Subject E chose to use Google for most of her queries.

Referring to the think-aloud protocol data in addition to screen-recording data proves revealing. Subjects who adopted the "single web resource approach" often appeared to be frustrated, probably due to the fact that it often took them a long time to locate satisfactory search results; occasionally they even failed to locate any satisfactory results at all. This is despite multiple attempts being made to vary search queries, sometimes using advanced features in their chosen search engine or web resource. Table 3 shows an example of this type of web-search style[vii].

INSERT TABLE 3 HERE

As seen in Table 3, Subject E chose to use Google exclusively as her web resource. She started by typing a ST term, "ropax" as an initial query in Google while verbalising "Is this a brand name?" She spent about 20 seconds browsing through her search engine result page (this can be observed from her mouse movements within the time frames being recorded). However, she decided not to click any of the search results. She then typed "ro-pax" in Google again. This time she clicked on a search result, which was Wikipedia. While browsing the Wikipedia page, she claimed that she understood the meaning of "ropax" but did not know how to render it in the TL. She then typed the query, "渡轮" [Pinyin: Dulun; literal translation: ferry] in Google and clicked the first search result, which is a Chinese encyclopaedia called, Hudong Baike (www.baike.com). On this page, she clicked two more links: "汽车渡轮" [Pinyin: Qiche Dulun; literal translation: car ferry] and "海上巴士" [Pinyin: Haishang Bashi; literal translation: sea bus] but still decided that there were no satisfactory TL equivalents. Therefore, she went back to Google to type another query, "可以运输车辆的渡轮" [Pinyin: Keyi Yunshu chelang de dulun; literal translation: a ferry that can transport cars]. According to her TAPs, she concluded that ropax basically means ferry. But, she still did not understand why the term, "ropax", was used in the ST, instead of "ferry". This is when she made a final search attempt, typing "ropax" again in Google. This time she clicked on the website, DFDSGroup.com. Having browsing the website for a few seconds, she showed her frustration by sighing and saying, "I am stuck here repeatedly" and finally decided that she would stop searching and tried to translate the term, "ropax" herself. It is clear that the reason Subject E struggled to find a TL equivalent for "ropax" is largely because she kept using Google rather than other types of web resources, since the TL equivalent for "ropax" can be easily found by using a bilingual dictionary on line or by using a Chinese search engine, such as Baidu, which was used by some of the other subjects in the present study.

On the other hand, the data also suggest that there is another type of web search behaviour where trainees chose to frequently use a variety of different types of web resources. Subject B and Subject F's cases typified this behaviour. In fact, upon closer examinations, it seems that

8

this type of subjects did not just have more varieties of web resources at their disposal but used different varieties of web resources to crosscheck a provisional search result. Table 4 illustrates this.

INSERT TABLE 4 HERE

As seen in Table 4, Subject B chose to pose the query, "STX 芬兰公司 [Pinyin: Fenlan gongsi; literal translation: Finland company]" at the timeframe 00.04.57. This query was posed because previously in Baidu, Subject B searched for the term, "STX Finland and Viking line". The search result in Baidu indicated that "STX Finland" should be translated as "STX 芬兰公司". Therefore, Subject B decided to use this result as a new query in Baidu to crosscheck its reliability and seek confirmation for this TT equivalent. This behaviour is reminiscent of Maniez's (2007) account of using the web as a mean to confirm terminology.

The two types of web-search behaviours have potential pedagogical implications which will be discussed in Section 3.4.

3.2 Queries

One of the questions this study sets out to answer is what search queries were made on the web, or, to put it another way, what constitutes Chinese trainee translators' web-search queries. The data show that queries predominantly consist of a source-text term and/or occasionally a target term, which range from a single lexical item (e.g. a word) to a phrase or sometimes even a sentence. However, there are also various other patterns found in queries[viii], which will be highlighted below.

1) A query can be a ST term and its corresponding (sometimes partially corresponding) TT equivalent.

For example, when searching for an abbreviated ST term, "LNG", Subject F typed the following two queries: "LNG 燃料船 [Pinyin: Ranliaochuan; literal translation: fuel ship]" and "LNG 燃料 [Pinyin: Ranliao; literal translation: fuel]" in Google. It appears that Subject F recognised that LNG is a type of fuel. Therefore, she decided to type queries that consist of both the original ST term and its partially understood TT equivalent.

2) A query can be a ST term and some question words (in the target language).

An example of this can be found in Subject A's query. After struggling to locate a TT equivalent for the ST abbreviated term, "GT", Subject A typed the following queries, "GT 是什么意思 [Pinyin: shi shenmo yisi; literal translation: what this means]" and "gt 是什么单位 [Pinyin: shi shenmo danwei; literal translation: what measuring unit this is]" in Baidu. This is similar to queries in the form of "what is …", often adopted in popular search engines, e.g. Google. Obviously, Subject A typed such questions in Chinese because Baidu is a Chinese search engine.

3) A query can be a ST term and the name of the TL

Another variation of the query is to simply type a ST term and the name of the TL. For example, Subject E typed the query, 'owner operator 中文 [Pinyin: Zhongwen; literal translation: Chinese]'. Again, this is a common advanced search engine strategy to search for a TL equivalent term, which is particularly useful for translators.

4) A query can be a provisionally translated sentence or title of the original ST

This is probably one of the most interesting queries. Subject D chose to translate the ST title tentatively, and used this tentatively translated ST title as a query. The ST title was "Viking builds world's largest LNG passenger ferry" and Subject D's query was "維京人建立世界第一个液化天然气渡轮" (Pinyin: Weijingren jianli shijie diyige yihua tianranqi dulun; literal translation: Viking builds world's first LNG ferry). This appeared to be a very effective query for locating a parallel text in the TL on the web. More importantly, the timing of this strategy

was also crucial. Instead of spending time searching for individual terms, Subject D chose to conduct this query in the first few minutes of her web-search process. This probably explains why she posed very few queries, in fact only 11 in total (see Table 2), compared to the rest of the subjects who entered between 24 and 88 queries. Upon closer examination, the majority of her potential queries were in fact solved by the parallel text that she located early on the web. This seems to demonstrate the advantage of locating parallel texts early on in the web search process.

3.3 *Other idiosyncratic web-search behaviour*

As mentioned in Section 3.1, Subject D and Subject F's idiosyncratic web-search behaviours are worth scrutinising further. They are presented below.

3.3.1 Subject D: using SMT as a dictionary

Subject D is the subject in the group who chose to use a SMT (i.e. Youdao Fanyi) almost exclusively as a web resource. In fact, as illustrated in Table 2, Subject D made 11 queries in total, of which 9 were via Youdao Fanyi. In addition, Subject D chose to use Youdao Fanyi as if it were a dictionary. As seen in Table 5, her queries in Youdao Fanyi were all ST terms.

INSERT TABLE 5 HERE

Subject D's reasons for choosing to use Youdao Fanyi as a dictionary may be uncertain, as unfortunately, she did not verbalise this in her think-aloud protocols. However, it is clear that she is aware of the existence of Youdao Fanyi's associated online dictionary (i.e. Youdao dictionary). In fact, at the beginning of her web-search process, she was found to deliberately ignore Youdao dictionary and went to Youdao Fanyi instead. Perhaps Subject D believed that Youdao Fanyi was more reliable than Youdao dictionary (which effectively is a collection of several paper-based dictionaries made available online). However, the fact that Subject D did not once use Youdao Fanyi to search any ST sentences indicates that, as a postgraduate trainee

translator, she was probably acutely aware of the quality issues associated with web-based SMT system for her language combination (English and Chinese) and yet she decided that Youdao Fanyi would be more helpful at lexical levels, particularly, locating TL terms, compared with online dictionaries.

3.3.2 Subject F: using Lingoes and Google

Compared to Subject D, Subject F also used a SMT system on the web, i.e. Google Translate. However, she used the SMT very differently. In fact, due to the pre-installed Lingoes package, Subject F was able to pose queries not just on Google Translate but on multiple web resources effortlessly. There are several levels of advantages to this. First of all, Subject F was able to pose 88 queries in total. This is a large number of queries made compared with other trainees in the study, even taking into account that Subject F did not conduct any web search prior to her translation (See Table 2). On the face of this, it shows that she was able to pose queries more efficiently and as a result, she had a lot more information and dictionary entries at her disposal within a very short period of time. This efficiency has profound ergonomical implications for her web-search process. It means that instead of typing queries in different web resources and constantly switching between her web browsers and her word processor, she was able to access different information on the web at a glance. In other words, the dynamics of the whole web-search process became very different compared with her fellow trainee translators.

For her fellow trainees, in order to search a ST term on the web, they might pose a query in an online dictionary first, and then decided whether the query result is satisfactory. If it was not, they could then decide whether they needed to pose the same query again in a different web resource or posed a slightly different query in the same web resource and this process might go on until they were reasonably satisfied with a search result or until they decide to abandon the search all together. Therefore, their cognitive web-search process appeared to be more linear and there was a clearer distinction between successive steps of this process.

In contrast, Subject F placed her cursor on a ST term and gained instant access to many different web-search results at once. This means that cognitively, her web-search process is not so much about searching for a TL equivalent one after another (at least initially) but more about deciphering the different varieties and larger volumes of information presented to her. More often than not, it is about selecting a (more) convincing TL equivalent among this pile of information rather than judging whether there is a need to pose another query on the web. It was only when Subject F decided that none of the information presented to her was satisfactory that she started to think about going outside the Lingoes package. She went to Google only intermittently to crosscheck or supplement her Lingoes search results.

INSERT FIGURE 2 HERE

Figure 2 An example of Subject F's use of Google Translate

Figure 2 shows an example of Subject F's use of SMT (i.e. Google Translate) where a whole sentence was selected as a query. Partly because unlike Subject D, Subject F conducted her web search while translating, it was possible to observe immediate effects and/or potential reasons of using SMT for translation. It was found that Subject F used Google Translate at two different levels and for different purposes. First of all, similar to Subject D, Google Translate was frequently used as if it were a dictionary, particularly when the other online dictionaries in Lingoes failed to locate TL equivalents for a particular ST term. When used as a dictionary, typical queries were at lexical level. Subject F was found to be using Google Translate either for comprehension purposes (e.g. checking or confirming the meaning of a term) and/or for expression purposes (i.e. looking for inspiration for a term) at this level. But, at clause and sentence levels, the purpose of using Google Translate was slightly broader. In other words, in addition to comprehension and expression, Subject F seemed to be using it as a quick fix to get the gist of a ST passage. This is probably very similar to how Google Translate is commonly used by the general public. What is more interesting is that towards the end of the translation process, she appeared to be relying on Google Translate a lot more frequently at the clause and sentence levels. There was also evidence to suggest that she was recycling some of the expressions or even sentence structures from Google Translate in her own translation at this stage. Effectively, she was partially doing post-editing towards the end of her translation

process. A possible reason for this is probably because she was feeling tired after an hour of web searching and translating and began to look for quick fixes for her to complete the translation.

Irrespective of potential reasons for this, this phenomena shows that the traditional boundary between translating and MT post-editing may be increasingly blurred. This is especially pertinent when one considers that an increasing number of translation memory (TM) systems have incorporated MT as an optional plug-in. This study demonstrates that the dichotomy of translating vs. post-editing may require a re-think, even when a (trainee) translator is not using such TM systems. Pedagogical implication of this will be discussed in the following.

3.4 Pedagogical implications

Potential pedagogical implications will be explored here in relation to Research Question 4. First of all, two types of web-search behaviours emerge from the data, as seen in Section 3.1. One type of trainees appears to succumb to the tendency of using one web resource only and was frustrated at times whereas the other type of trainees used more varieties of web resources and often use them to cross-check and confirm search results. As a result, the second type of trainee showed less frustration in their web search. There are some obvious pedagogical implications to this; that is, it is important for trainees to understand the drawback of predominantly using a single and habitual web resource. Given that most web browsers have their default search engines and automatic memories of previously used web pages, it might be all too easy for trainees to fall into this trap of using the same web pages or habitual resources repeatedly. Rather, students should be encouraged to use a combination of different types of web resources more flexibly because multiple web resources might not just maximize chances of locating suitable translation equivalent but also potentially validate and cross-check located translation equivalent when necessary and consequently resulted in successful web search.

Another important finding that may have pedagogical implication is related to trainees' increasing use of online SMT systems. As mentioned earlier, the boundary between translating

and post-editing appears to be increasingly blurred because the use of online SMT appears to be intertwining with or even considered to be an integral part of trainees' problem solving behaviours. From a pedagogical point of view, it is, therefore, out-of-date to assume or even advise trainees not to trust such SMT systems. In fact, just as Garcia (2011) has predicted, post-editing may well be the way forward. This situation is reminiscent of the debate years ago when Wikipedia first became available on the web. At first, it was seen by many as an unreliable source, reasons being that surely an encyclopaedia that could be contributed and amended by anyone at any time could not be trusted, particularly from translators' points of view. Trainees were advised not trust such a volatile web resource. However, overtime, Wikipedia has both been substantially improved and immensely expanded. Nowadays, Wikipedia has become one of the most popular online encyclopaedias. The same is true for online SMT, e.g. Google Translate. The reality is, as shown in the present study, trainees have already embraced such tools on the web. But, currently very little is known about how various tools, including online SMT systems (e.g. Google Translate), hybrid TM and MT systems (e.g. Google Translators' Toolkit or Wordfast Anywhere) or indeed a combination of these two and other web resources may have shaped and conditioned trainee or possibly professional translators' working practice and their cognitive process. While it is important to incorporate SMT into translation curriculum, just as Kenny and Doherty (2014a, 2014b) has indicated, it will be difficult to do so without further lines of research into this.

4. Conclusion

This study set out to explore Chinese trainee translators' web-search behaviours. In terms of the types of web resources used, it was found that Chinese trainee translators largely use similar types of web resources, such as search engines, online dictionaries, online encyclopaedias and online MT systems, as indicated by previous researchers. In terms of the frequency of use for each type of web resource, some Chinese trainees were found to use online dictionaries mostly. The predominant use of online dictionaries is an interesting phenomenon, since it demonstrates that dictionaries (especially bilingual ones) are still a primary consultation source for Chinese trainee translators, even though it is web-based rather than paper-based ones that they use. It will be of great interest to find out whether the same applies to other translators, e.g. veteran professionals and for translators working in other language combinations.

Search queries form another important part of the study. It was found that Chinese trainees typically use source-text terms as their queries. However, their queries could also consist of TT terms associated with particular ST terms. At other times, a query consisted of different varieties of both a ST term and a TT term. Most interestingly, one trainee chose to use a tentatively translated ST title as a query. This appeared to be an effective strategy for locating parallel texts at an early stage of web search.

In terms of idiosyncratic behaviours, data show that an online SMT system was used as a dictionary. Yet, there was evidence to show that an online SMT system was used more extensively, particularly towards the end of the translation process. Partly because of the nature of the online translation and dictionary package being used (i.e. Lingoes), this particular trainee's cognitive web-search behaviours appeared to be (ergonomically) more efficient when compared with other trainees. Partly because of the same reason, this trainee was found to recycle the SMT output in her translation. In other words, she was effectively doing post-editing and producing her translation at the same time. This finding points to an interesting direction for future research in terms of how such online SMT or even the new hybrid TM and TM system may have moulded translators' working practice.

Apart from these findings, the present study further demonstrates the feasibility of adopting screen-recording (see Göpferich & Jääskeläinen 2009, 173) with think-aloud protocols in investigating web-search behaviour, even though it will also be beneficial to add post-task interviews so that data can be triangulated and further validated. Similarly, given the limited number of subject population in the present study, its findings may not apply to all translators. Therefore, it will be beneficial for future research to scale up the subject population and language combinations.

References

Buendía-Castro, Miriam and Clara Inés López Rodríguez. 2013. "The Web for Corpus and the Web as Corpus in Translator Training." *International Journal of Lexicography* 23(3):54-71.

Chesterman, Andrew. 2009. "The Name and Nature of Translator Studies." *Hermes - Journal of Language and Communication Studies* 42:13-22.

Daems, Joke; Carl, Michael; Vandepitte, Sonia; Hartsuiker, Robert; and Lieve Macken. 2016. "The Effectiveness of Consulting External Resources During Translation and Post-Editing of General Text Types." In *New Directions in Empirical Translation Process Research: Exploring the CRITT TPR-DB*, edited by Carl, Michael; Bangalore, Srinivas; and Moritz Schaeffer, 111-134. New York, Dordrecht, London: Springer International Publishing, Switzerland.

Ehrensberger-Dow, Maureen; Englund Dimitrova, Birgitta; Hubscher-Davidson, Séverine; and Ulf Norberg (eds). 2015. *Describing Cognitive Processes in Translation: Acts and Events*. Amsterdam/Philadelphia, John Benjamins.

Enríquez Raído, Vanessa. 2011. "Developing Web Searching Skills in Translator Training." *Redit: Revista Electrónica de Didáctica de la Traducción y la Interpretación* 6:60–80.

Enríquez Raído, Vanessa. 2014. *Translation and Web Searching*. New York and London: Routledge.

Garcia, Ignacio. 2009. "Beyond Translation Memory: Computers and the Professional Translator." *JoSTrans- The Journal of Specialised Translation* 12: 199-214. http://www.jostrans.org/issue12/art_garcia.php

Garcia, Ignacio. 2011. "Translating by Post-editing: Is It the Way Forward?." *Machine Translation* 25(3): 217-237.

Göpferich, Susanne and Riitta Jääskeläinen. 2009. "Process Research into the Development of Translation Competence: Where Are We, and Where Do We Need to Go?" *Across Languages and Cultures* 10(2):169–191.

Gouadec, Daniel 2007. *Translation as a Profession*. Amsterdam: John Benjamins

Harman, Nicky. 2002. *Beyond Paper Dictionaries: Mining the Web for Technical Terminology in Chinese. Paper Prepared for the Translation Teacher's Certificate of*

the Consortium for the Training of Translation Teachers.
http://isg.urv.es/cttt/cttt/research/harman2.pdf. Accessed August 2, 2015.

Jääskeläinen, Riita. 1989. "Translation Assignment in Professional vs. Non-Professional
Translation: A Think-Aloud Study." In *The Translation Process*, edited by T C
Seguinot, 87–98. Ontario: York University Press.

Jansen, Bernard and Udo Pooch. 2001. "A Review of Web Searching Studies and a
Framework for Future Research." *Journal of American Society for Information Science
and Technology* 52(3):235–246.

Kenny, Dorothy and Stephen Doherty. 2014. "Statistical Machine Translation in the
Translation Curriculum: Overcoming Obstacles and Empowering Translators." *The
Interperter and Translator Trainer* 8(2): 276-294.

Kenny, Dorothy and Stephen Doherty. 2014. "The Design and Evaluation of a Statistical
Machine Translation Syllabus for Translation students." *The Interperter and Translator
Trainer* 8(2): 295-315.

Krings, Hans P. 1986. "Translation Problems and Translation Strategies of Advanced German
Learners of French (L2)." In *Interlingual and Intercultural Communication*, edited by J
House and S Blum-Kulka, 263–276. Tubingen: Gunter Narr Verlag.

Künzli, Alexander. 2001. "Experts versus novices: l'utilisation de sources d'information
pendant le processus de traduction." *Meta: Journal des traducteurs* 46(3):507–523.

Levy, J. 1967/2000. "Translation as a Decision Making Process." In *The Translation Studies
Reader*, edited by Lawrance Venuti, 148-159. London and New York: Routledge.

Livbjerg, Inge and Inger M. Mees. 2003. "Patterns of Dictionary Use in Non-Domain-
Specific Translation." In *Triangulating Translation: Perspectives in Process Oriented
Research*, edited by Fábio Alves, 123–136. Amsterdam/Philadelphia: John Benjamins.

Maniez, François. 2007. "Using the Web and Computer Corpora as Language Resources for
the Translation of Complex Noun Phrases in Medical Research Articles." In *Panace*
8(26):162-167.

Massey, Gary and Maureen Ehrensberger-Dow. 2011. "Technical and Instrumental Competence in the Translator's Workplace: Using Process Research to Identify Educational and Ergonomic Needs." *ILCEA* 14:2–12.

Mayor Serrano, Blanca. 2009. "Recursos en línea relacionados con el ámbito marítimo y naval [Online resources in the naval and maritime field]." *The Translation Journal* 13(1).

Mikhailov, Mikhail. 2015. "Minor Language, Major Challenges: the Result of a Survey into the IT Competences of Finnish Translators." *JoSTrans- The Journal of Specialised Translation* 24:89-111. http://www.jostrans.org/issue24/art_mikhailov.php

Pinto, María and Dora Sales. 2007. "A Research Case Study for User-Centred Information Literacy Instruction: Information Behaviour of Translation Trainees." *Journal of Information Science* 33(5):531-550.

Pinto, María and Dora Sales. 2008a. "Towards User-Centred Information Literacy Instruction in Translation." *The Interpreter and Translator Trainer* 2(1):47–74.

Pinto, María and Dora Sales. 2008b. "INFOLITRANS: A Model for the Development of Information Competence for Translators." *Journal of Documentation* 64(3):413-437

Pym, Anthony 2011. "What technology does to translating." Translation and Interpreting 3(1): 1-9

Wakabayashi, Judy. 2003. "Think Aloud as a Pedagogical Tool." In *Beyond the Ivory Tower: Rethinking Translation Pedagogy*, edited by B J Baer and G S Koby, 61–82. Amsterdam: John Benjamins.

Xu, Mianjun and Caiwen Wang. 2011. "Translation Students' Use and Evaluation of Online Resources for Chinese-English Translation at the Word Level." *Translation and Interpreting Studies* 6(1):62–86.

Zanettin, Federico. 2002. "DIY Corpora: the www and the Translator." In *Training the Language Services Provider for the New Millennium*, edited by Belinda Maia, Johann Haller, Margherita Ulrych, 239-248. Porto: Universidade do Porto. Faculdade de Letras.

Zheng, Binghan. 2014. "The Role of Consultation Sources Revisited: An Empirical Study of English–Chinese Translation." *Perspectives* 22(1):113–135.

Appendix A

Task instructions

In this task, you are asked to think aloud and record your on-screen activities using BBFlashBack Express. However, before you start to translate, you are asked to search the World Wide Web to support your translation of this semi-specialised text. In other words, you will be recording your web search and think-aloud protocols using BBFlashBack Express. Please remember to select the **"RECORD SOUND"** button before you start recording.

Your source text starts here…

Viking builds world's first large LNG passenger ferry

When Finnish owner-operators STX Finland and Viking Line sought a technical partner to help design and classify a new ropax ship fuelled by LNG, they approached Lloyd's Register with its renowned expertise in LNG-as-a-fuel.

The result will be the world's first large ferry to use liquid natural gas as its main source of power. The 56,850 gt ferry will be built at Finland's STX Turku shipyard, one of Europe's largest yards and the source of the world's biggest cruise liners, *Allure of the Seas* and *Oasis of the Seas*. One of the options in the newbuild contract is for a sister LNG ferry to be built.

The huge ferry which is due to be delivered at the start of 2013 will be the world's most environmentally-friendly passenger ship with minimal CO_2 and virtually nil SO_x and NO_x emissions. Wave formation and noise generation on the 210m-long vessel will also be negligible.

With a maximum speed of almost 22 knots, the ropax will be able to carry 2,800 passengers and will have a hoistable car-deck with a vehicle capacity of 1,100 lane-metres (lm) and 1,300 lm of truck capacity. It will operate between Turku and Sweden's Stockholm and is being designed to cope with the comparatively sensitive and shallow waters of the Finnish archipelago.

"LNG tends to suit ferries better than other types of vessel as they follow fixed routes between ports with LNG terminals in the neighbourhood which makes them easy to bunker. The same bunkering facilities aren't yet available for other ships including cruise ships," said project manager Matti Niskala, LR's Finnish Marine Country Business Manager.

However with the innovative groundwork that is being achieved on the current project, Niskala and his team believe other vessels should be able to make the switch to LNG in five to 10 years' time.

"The ropax will be the first newbuild to comply with our provisional rules for LNG propulsion although we've overseen and applied the same set of rules to a number of other vessels, notably the *Accolade* bulk carrier which is the world's first LNG-fuelled ship. Since the STX-Viking contract was signed, we have also provided the STX technical team with assistance on issues such as developing LNG storage facilities and process piping," said Niskala.

"We carried out a detailed risk analysis for Viking Line on the bunkering process to identify and minimise risks associated with the movement of the bunker barge and ship within the confines of the port, the risks associated with the simultaneous loading of passengers, cars, lorries and LNG, and to help ensure compatibility between bunker barge capacity and the ship's systems."

The ropax ferry project poses many technical challenges for both owner and builder. "As well as ensuring the safety of the LNG system, a key issue will be the integration of both LNG and oil fuel installations and compliance with the new IMO 'safe return to port' requirements, a procedure on which Lloyd's Register is a world leader," said Niskala.

---

[i] Enríquez Raído 's two professional translators only completed Task 1 in her pilot study.

[ii] 'Query' will be defined in Section 2 and discussed in Section 3.2.

[iii] These trainee subjects, whose age ranges from 23 to 30, were all female and had language-related undergraduate degrees prior to studying for their MA in Translation Studies. Their English language proficiency was IELTS 6.5 or above. These subjects were considered to represent a stereotypical subject population of the above mentioned degree, although their participation was on a volunteering basis.

[iv] The actual instructions and source text given to trainees can be found in Appendix A.

[v] Due to the sheer size of these files, subjects were asked to submit their recording to the researcher via DropBox™. These files were submitted at the end of semester two during trainees' year-long postgraduate degree.

[vi] The number of uses for each web resource is counted on the basis of a primary query, rather than secondary search results. For example, when a subject typed a query, "knots" in Google and then clicked on a number of Google search results, which may include a link to Wikipedia, an online dictionary or other webpages, etc. This is still counted as one use of Google.

[vii] Subjects' think-aloud protocols (TAP) were originally in Mandarin Chinese. They were transcribed and translated into English and presented in Italic by me.

[viii] This study did not claim to cover all types of queries, given that it is based on a relatively small subject population. However, it is important to present various query patterns based on the data available in the present study since there are very few studies looking into queries patterns in translation web search.